Journal of Biomolecular Techniques • Volume 34(1); 2023 Apr

Article Watch: April, 2023

Clive Slaughter[1]

[1]**AU-UGA Medical Partnership**

Association of Biomolecular Resource Facilities

Published on: Mar 27, 2023

DOI: https://doi.org/10.7171/3fc1f5fe.d658481f

ABSTRACT

This column highlights recently published articles that are of interest to the readership of this publication. We encourage ABRF members to forward information on articles they feel are important and useful to Clive Slaughter, AU-UGA Medical Partnership, 1425 Prince Avenue, Athens GA 30606. Tel: (706) 713-2216; Fax: (706) 713-2221; Email: cslaught@uga.edu or to any member of the editorial board. Article summaries reflect the reviewer's opinions and not necessarily those of the Association.

NUCLEIC ACID SEQUENCING

Begik O, Diensthuber G, Liu H, Delgado-Tejedor A, Kontur C, Niazi A M, Valen E, Giraldez A J, Beaudoin J-D, Mattick J S, Novoa E M. Nano3P-seq: transcriptome-wide analysis of gene expression and tail dynamics using end-capture nanopore cDNA sequencing. *Nature Methods* 20;2023:75-85.

Polyadenylation of RNA affects the stability and translational efficiency of RNA molecules. Begik *et al*. describe a new method for investigation of polyadenylation that is based on long-read nanopore cDNA sequencing. The method uses template switching to initiate reverse transcription to avoid requirements for 3' end adapter ligation steps, PCR amplification, and second-strand cDNA synthesis, thereby minimizing biases introduced in library preparation. The method enables sequencing of RNA molecules from the 3' end, regardless of the presence or absence of a polyA tail, or its length, or the presence of non-A bases within it. These characteristics permit simultaneous quantification of both coding and non-coding RNA abundance, as well as measurement of polyA tail lengths of individual RNA molecules and the length distribution within individual isoforms. The methodology therefore provides the means to investigate how cells use dynamic regulation of polyadenylation in the post-transcriptional control of gene expression. The authors use it to investigate the dynamics of polyadenylation during the transition from translation of maternal RNA to zygote RNA after fertilization in zebrafish development.

Amman F, Markt R, Endler L, Hupfauf S, Agerer B, Schedl A, Richter L, Zechmeister M, Bicher M, Heiler G, Triska P, Thornton M, Penz T, Senekowitsch M, Laine J, Keszei Z, Klimek P, Nägele F, Mayr M, Daleiden B, Steinlechner M, Niederstätter H, Heidinger P, Rauch W, Scheffknecht C, Vogl G, Weichlinger G, Wagner A O, Slipko K, Masseron A, Radu E, Allerberger F, Popper N, Bock C, Schmid D, Oberacher H, Kreuzinger N, Insam H, Bergthaler A. Viral variant-resolved wastewater surveillance of SARS-CoV-2 at national scale. *Nature Biotechnology* 40;2022:1814-1822.

Amman *et al.* document the use of wastewater-based surveillance of SARS-CoV-2 infection on a national scale. Their study is conducted in Austria, a central-European country with a comprehensive SARS-CoV-2 epidemiologic surveillance program that provides complementary information about clinical variants, and has a population with 93% of individuals connected to sewer infrastructure, which allows broad-based sampling of virus production from wastewater. The authors perform high-throughput sequencing on 3,413 wastewater

samples derived from 94 municipalities covering >59% of the country's population to deduce the spatiotemporal distribution of pre-defined viral variants. The results are validated by comparison with the epidemiologic records of >311,000 individual cases. Relative abundance and reproduction numbers of variants are calculated, and *de novo* detection of emerging variants is demonstrated. This approach to surveillance will gain importance as sampling bias and economic constraints on standard epidemiologic approaches increase. Nevertheless, interpretation of wastewater data will continue to rely upon information about the amount and duration of viral shedding in feces from patients with changing immune status, and information about altered properties of evolving viral variants.

Kjær K H, Winther Pedersen M, De Sanctis B, De Cahsan B, Korneliussen T S, Michelsen C S, Sand K K, Jelavić S, Ruter A H, Schmidt A M A, Kjeldsen K K, Tesakov A S, Snowball I, Gosse J C, Alsos I G, Wang Y, Dockter C, Rasmussen M, Jørgensen M E, Skadhauge B, Prohaska A, Kristensen J Å, Bjerager M, Allentoft M E, Coissac E, Alsos I G, Coissac E, Rouillard A, Simakova A, Fernandez-Guerra A, Bowler C, Macias-Fauria M, Vinner L, Welch J J, Hidy A J, Sikora M, Collins M J, Durbin R, Larsen N K, Willerslev E, Phylonorway C. A 2-million-year-old ecosystem in Greenland uncovered by environmental DNA. *Nature* 612;2022:283-291.

This study demonstrates that metagenomic sequencing of environmental DNA may be used to reconstruct an ancient paleoenvironment from the Early Pleistocene period. The DNA sampled here is derived from sediments of terrestrial origin washed into an estuary, and sediments deposited in a deep-water marine environment, situated in northern Greenland. The area is now polar desert, but 2 million years ago it was interglacial boreal forest. Identified terrestrial animals include reindeer, caribou and mastodon, rodents and geese – all herbivores, but no carnivores, probably because of the greater biomass of herbivorous species. Plants include poplar, birch and thuja trees, and various Arctic and boreal shrubs and herbs. Plant species predominate over metazoans, probably for similar reasons related to biomass. Marine species include horseshoe crab and green algae. The study illustrates the capability of ancient environmental metagenomic sequencing to provide information for reconstruction of ancient environments, study of the phylogeny, dating of ancient lineages of plants and animals, and delineation of the geographic distribution of ancient genera.

Kennedy K M, De Goffau M C, Perez-Muñoz M E, Arrieta M-C, Bäckhed F, Bork P, Braun T, Bushman F D, Dore J, De Vos W M, Earl A M, Eisen J A, Elovitz M A, Ganal-Vonarburg S C, Gänzle M G, Garrett W S, Hall L J, Hornef M W, Huttenhower C, Konnikova L, Lebeer S, Macpherson A J, Massey R C, Mchardy A C, Koren O, Lawley T D, Ley R E, O'mahony L, O'toole P W, Pamer E G, Parkhill J, Raes J, Rattei T, Salonen A, Segal E, Segata N, Shanahan F, Sloboda D M, Smith G C S, Sokol H, Spector T D, Surette M G, Tannock G W, Walker A W, Yassour M, Walter J. Questioning the fetal microbiome illustrates pitfalls of low-biomass microbial studies. *Nature* 613;2023:639-649.

This *Perspective* addresses a general issue in microbial studies concerning interpretation of sequence-based data from samples of low biomass, in which some or all of the signals may be derived from background contamination. The impact of contamination in studies of high biomass samples, *e.g.* feces, dental plaque, wastewater, and soil, is likely to be very low. In studies of low biomass samples, *e.g.* skin swabs, nasal tract

swabs, breast milk, respiratory tract, and tissue biopsy samples, however, sample contamination may become progressively more prominent as biomass content of the samples themselves decreases. The authors suggest mitigation strategies that include (but are not limited to) sample concentration to increase biomass before processing, processing in a rigorously clean-room environment, avoidance of kit-based methods in favor of decontaminated reagents and surfaces, the use of negative controls, and inclusion of "mock communities" represented by DNA of known composition to assess by serial dilution the biomass level at which contamination becomes a dominant feature of the sequencing results. In studies of samples in which the existence of microorganisms has not been consistently established, *e.g.* placental and fetal tissues, amniotic fluid, meconium, and biopsy samples of cerebrospinal fluid, blood, or solid tissues from healthy individuals, the authors additionally advocate the use of quantitative methods such as qPCR or direct microscopic examination following Gram staining or FISH to determine whether microorganisms are indeed present prior to sequence analysis. The presence of microorganisms identified by sequencing should be confirmed by replication with independent samples, diverse processing techniques, and different batches of any kits that are used. Sequencing should be pursued to high taxonomic resolution to enable subsequent confirmation of conclusions by species-specific qPCR or FISH. The authors critically evaluate studies of samples from the prenatal uterine environment in light of these considerations, and show how interpretation of the data have a fundamental impact on our understanding of how the immune system develops.

MASS SPECTROMETRY

Cooper-Shepherd D A, Wildgoose J, Kozlov B, Johnson W J, Tyldesley-Worster R, Palmer M E, Hoyes J B, Mccullagh M, Jones E, Tonge R, Marsden-Edwards E, Nixon P, Verenchikov A, Langridge J I. Novel hybrid quadrupole-multireflecting time-of-flight mass spectrometry system. *Journal of the American Society for Mass Spectrometry* 34;2023:264-272.

Time-of-flight mass spectrometry (TOF-MS) is capable of high acquisition speeds up to several hundreds of spectra per second. The methodology is therefore suitable for profiling fast separations encountered in ultra-high-performance liquid chromatography, gas chromatography, capillary electrophoresis, and ion mobility spectrometry, and for acquisition of large mass spectrometric images at high spatial resolution. The present paper describes increases in mass resolution and mass accuracy achieved with a multi-reflecting TOF mass analyzer system that nonetheless preserves speed of data acquisition. Ions pass through a quadrupole mass filter, followed by a segmented, gas-filled, quadrupole collision cell, then encounter an inclined double-orthogonal accelerator/deflector that pushes them into the TOF sector. There, they traverse a multi-reflecting TOF analyzer that uses gridless mirrors, which eliminate ion losses associated with gridded mirrors. Ion packet divergence is controlled by an array of periodic focusing lenses located midway between the mirror sets. The system performs 46 reflections to accomplish a path-length of ~48 m (flight time of ~1.3 ms for *m/z* 1000), producing a resolving power of >200,000 fwhm and sub-ppm mass accuracy. Instead of waiting for all ions to clear the TOF sector between pulses, Encoded Frequent Pushing rather than a fixed pulse period permits decoding of the products of individual pulses to improve the sampling duty cycle to >10%, yielding 128-fold

improvement in sensitivity and extension of the lower end of the dynamic range for which sub-ppm accuracy is achievable. The utility of the system is demonstrated in LC-MS and mass spectral imaging applications in which high resolution and high mass accuracy strongly enhance analyte identification. The methodology provides performance typically associated with FT-ICR systems, but without incurring long transient times to achieve high resolving power. A commercial implementation of this technology is available in the SELECT SERIES MRT mass spectrometry platform from Waters Corporation, Milford, MA.

Salome A Z, Lee K W, Grant T, Westphall M S, Coon J J. Matrix-landing mass spectrometry for electron microscopy imaging of native protein complexes. *Analytical Chemistry* 94;2022:17616-17624.

A large body of evidence indicates that protein complexes desorbed into the gas phase as cations for the purposes of mass spectrometric analysis retain many features of their liquid-phase structure and binding characteristics. Prior work in the authors' laboratory shows that protein complexes subjected to mass separation in the gas phase may be collected by soft-landing on a carbon grid, and may then be subjected to transmission electron microscopy (transmission EM) to reveal information about their 3-D structure. The key innovation in this earlier study was the use of a thin film of glycerol as a landing matrix. The present work extends and improves this methodology by investigating the utility of alternative soft-landing conditions. The authors find that poly(propylene) glycerol (2700 Da) performs better than glycerol as a landing matrix. They present negative stain transmission electron microscope images of GreEL, archeal 20 S proteasome, and β-galactosidase complexes, and reconstruct the 3-D structures of these complexes to ∼20 Å resolution. The mechanism by which the soft landing conditions preserve the native structure of the proteins is not yet fully determined, but the results validate the methodology as a means for coupling mass spectrometry with transmission electron microscopy. The data encourage optimism that further development of the methodology will enable coupling with cryo-EM imaging at even higher resolution to permit acquisition of structural information.

IMAGING

Zheng L, Liu N, Gao X, Zhu W, Liu K, Wu C, Yan R, Zhang J, Gao X, Yao Y, Deng B, Xu J, Lu Y, Liu Z, Li M, Wei X, Wang H-W, Peng H. Uniform thin ice on ultraflat graphene for high-resolution cryo-EM. *Nature Methods* 20;2023:123-130.

For minimization of background noise, cryo-EM imaging requires presentation of samples in a thin liquid film from which a thin, homogeneous sheet of vitrified ice is formed. Unfortunately, inhomogeneity and roughness of the underlying support leads to inhomogeneity of the vitreous ice layer, which degrades image resolution. Zheng *et al.* prepare ice sheets of enhanced uniformity by means of a supporting film of improved flatness. They replace copper foil, on which a graphene supporting film is normally deposited, with an ultra-flat Cu(III)/sapphire wafer as a growth substrate. The resulting ultra-flat graphene layers have improved mechanical strength and display height variations reduced by ∼10x. Image quality of protein complexes is

improved, and the authors demonstrate the reconstruction of a 3-D structure for hemoglobin (64 kDa) at a resolution of 3.5 Å, for α-fetoprotein (67 kDa) at 2.6 Å, and for streptavidin (52 kDa) at 2.2 Å. The authors envision general application of the methodology for atomic-resolution EM imaging and drug design.

Hu H, Huang H, Li M, Gao X, Yin L, Qi R, Wu R S, Chen X, Ma Y, Shi K, Li C, Maus T M, Huang B, Lu C, Lin M, Zhou S, Lou Z, Gu Y, Chen Y, Lei Y, Wang X, Wang R, Yue W, Yang X, Bian Y, Mu J, Park G, Xiang S, Cai S, Corey P W, Wang J, Xu S. A wearable cardiac ultrasound imager. *Nature* 613;2023:667-675.

Echocardiography – ultrasound imaging of the heart – is routinely used for clinical assessment of cardiac anatomy, and for measurement of functional parameters such as stroke volume and ejection fraction. In current procedures, cardiac output can be measured before and after exercise, but continuous monitoring during exercise is not practicable. The present paper describes proof-of-principle studies characterizing a wearable ultrasonic patch that may circumvent this limitation. It encapsulates phased arrays of piezoelectric transducers and liquid metal composite electrodes. Its flexible mechanical characteristics allow the device to maintain close contact with the skin over a large area while being worn. The transducer elements adopt a non-planar distribution as they mold to the human chest. The phase distortion that results from non-planarity is corrected with the use of a three-dimensional scanner for collection of data on chest curvature. The authors confirm image quality by evaluating axial, lateral and elevational spatial resolution, signal-to-noise ratio, axial and lateral location accuracies, dynamic range, and contrast-to-noise ratio. They apply deep learning neural network models for image processing to extract key clinical information automatically from the continuous stream of images. Next steps will include accommodation of dynamic changes in chest curvature to improve compensation of phase distortion and spatial resolution presently provided on a static basis by the scanner, and application of the neural network to subjects outside the training dataset. The present results nevertheless encourage optimism that the technology can be extended to imaging of other deep tissues and that it may be deployed for the care of outpatient and athletic patient populations.

MACROMOLECULAR SYNTHESIS & SYNTHETIC BIOLOGY

Stevens A J, Harris A R, Gerdts J, Kim K H, Trentesaux C, Ramirez J T, Mckeithan W L, Fattahi F, Klein O D, Fletcher D A, Lim W A. Programming multicellular assembly with synthetic cell adhesion molecules. *Nature* 614;2023:144-152.

Binding interactions between cells, and between cells and extracellular matrix components, are mediated by cell adhesion molecules (CAMs). CAMs are transmembrane proteins. Extracellular domains mediate adhesion, and intracellular domains interact with intracellular components such as the cytoskeleton to control cellular responses to the binding interaction. Stevens *et al.* explore the possibility that cellular expression of synthetic CAMs with recombined extracellular and intracellular domains might generate new kinds of cellular connectivity. Their approach is to express in L929 mouse fibroblasts (a cell line with low endogenous adhesion) CAMs whose extracellular domain is either green fluorescent protein (GFP) or a GFP nanobody

(anti-GFP) with which it can bind with defined affinity. The intracellular domains are drawn from a panel of CAMs with diverse properties. They include cadherins and integrins. When the cells expressing recombined CAMs interact with one another, or interact with a GFP-derivatized surface, the authors characterize the interaction phenotype by assessing the interface area, and the degree of GFP enrichment at the interface. They show that the choice of intracellular domains largely determines distinct mechanical and morphological characteristics of cellular interaction, and, when different intracellular domains are present on the two sides of the interface, the balance of their properties determines the interface morphology. By choosing extracellular domains of different specificity and coupling them with intracellular domains of different functionality, the authors program customized multicellular assemblies. Their results indicate that synthetic CAMs will provide powerful means for study of multicellular organization, for remodeling or repairing tissues, and for modulating cellular interactions.

METABOLOMICS & PROTEOMICS

Muehlbauer L K, Jen A, Zhu Y, He Y, Shishkova E, Overmyer K A, Coon J J. Rapid multi-omics sample preparation for mass spectrometry. *Analytical Chemistry* 95;2023:659-667.

The authors of this paper describe a simple, rapid, consolidated protocol for extraction of metabolites, lipids and proteins from a single sample suitable for LC-MS analysis. Robust, reproducible protocols for this purpose are available. They generally involve splitting samples for extraction of the different classes of molecule, or else involve some combination of metabolite and lipid extraction with a biphasic organic solvent system to separate molecules on the basis of polarity, and precipitation of proteins followed by re-solubilization for trypsin digestion. Such protocols are complex, lengthy, and prone to poor or biased analyte recovery. Muehlbauer *et al.* instead use monophasic solvent extraction with *n*-butanol in the presence of hydrophilic but unfunctionalized magnetic beads. High concentration of organic solvent induces hydrophilic interaction between protein and beads. Unbound metabolites and lipids are removed for further analysis, and the protein-bound beads are rinsed, the protein digested, and desalted. Digestion is accelerated by heating to 70˚C for completion within 1 h. The entire procedure is complete in 3 h.

FUNCTIONAL GENOMICS

Loyfer N, Magenheim J, Peretz A, Cann G, Bredno J, Klochendler A, Fox-Fisher I, Shabi-Porat S, Hecht M, Pelet T, Moss J, Drawshy Z, Amini H, Moradi P, Nagaraju S, Bauman D, Shveiky D, Porat S, Dior U, Rivkin G, Or O, Hirshoren N, Carmon E, Pikarsky A, Khalaileh A, Zamir G, Grinbaum R, Abu Gazala M, Mizrahi I, Shussman N, Korach A, Wald O, Izhar U, Erez E, Yutkin V, Samet Y, Rotnemer Golinkin D, Spalding K L, Druid H, Arner P, Shapiro A M J, Grompe M, Aravanis A, Venn O, Jamshidi A, Shemer R, Dor Y, Glaser B, Kaplan T. A DNA methylation atlas of normal human cell types. *Nature* 613;2023:355-364.

Although DNA methylation is linked closely to gene expression, DNA accessibility and chromatin packaging, the majority of existing surveys of DNA methylation in mammalian tissues are limited in scope through such

factors as the number of methylation sites sampled, or reliance on bulk tissue samples or cultured cells. In the present study, deep genome-wide paired-end bisulfite sequencing is performed to acquire 150-bp-long reads at an average sequencing depth of 30x (6.62x or greater) using populations of cells purified from 39 freshly dissociated adult, healthy tissues by fluorescent-activated cell sorting (FACS). Cell type differences in methylation patterns are identified on a genome-wide scale by identifying in each cell type blocks of homogeneously methylated CpG sites. The resulting DNA methylation atlas illuminates the biological impact of methylation and provides a resource for the design of further studies. Cells belonging to the same cell type are strikingly similar in methylation pattern (>99.5% identical). Different cell types have different unmethylated genomic regions that are mostly unique to one cell type. Methylation patterns of different cell types cluster in ways that more closely reflect developmental origin than gene expression pattern. The distinguishing unmethylated blocks occur predominantly in regions of accessible DNA, particularly at binding sites for tissue-specific transcriptional regulators. Uniquely hypermethylated blocks, which are much less common, are interestingly enriched in CTCF binding sites, suggesting a role for DNA methylation in attenuating CTCF binding and thereby modulating local 3-D organization of chromatin. The authors anticipate diverse uses for their methylome atlas, including identification of genetic variants predisposing to disease and identification of biomarkers for use in liquid biopsies.

DESIGN & DEVELOPMENT OF THERAPEUTICS

Lebek S, Chemello F, Caravia X M, Tan W, Li H, Chen K, Xu L, Liu N, Bassel-Duby R, Olson E N. Ablation of CaMKIIδ oxidation by CRISPR-Cas9 base editing as a therapy for cardiac disease. *Science* 379;2023:179-185.

A broad range of cardiac diseases, including ischemia/reperfusion injury, heart failure, cardiac hypertrophy and arrhythmias, has been linked to inflammation, cell death and fibrosis that are mediated by activation of Ca^{2+}/calmodulin-dependent protein kinase IIδ (CaMKIIδ). Activation of this enzyme can result from oxidation of 2 methionine residues, Met281 and Met282, located in the regulatory domain of CaMKIIδ. Oxidation of these residues activates the enzyme by preventing association between the enzyme's catalytic domain and its regulatory domain. Replacement of these methionine residues by valines in the germline of knock-in mice prevents oxidative activation and confers cardioprotection. Lebek *et al.* here ablate the oxidative activation site of CaMKIIδ by CRISPR-Cas9 adenine base editing in human induced pluripotent stem cells, and show that protection is thus conferred against ischemia/reperfusion injury. The authors proceed to show that delivery of base editing components packaged in adeno-associated virus serotype-9 (AAV9) at the time of experimental ischemia/reperfusion injury in mice preserves cardiac function. The authors propose that, following myocardial infarction, delivery of CaMKIIδ editing components *via* catheter could be performed in conjunction with standard-of-care revascularization in human patients. Their work opens prospects for a major extension of CRISPR-Cas9 gene editing therapy from the correction of relatively rare monogenic mutations to the interruption of a signaling pathway that mediates cardiac disease of high prevalence.